On mixture approximations for differential equations with distributed time delays

Tobias K. S. Ritschel

Assistant Professor
Department of Applied Mathematics and Computer Science,
Technical University of Denmark

August, 2025

Outline

1. What are mixture approximations?
   - Three different types of mixture approximations
   - Convergence
   - Practical implementation
   - What can you do with them?
2. Mixture approximations for delay differential equations (DDEs)
   - Examples of DDEs with distributed time delays
   - The linear chain trick (DDEs $\rightarrow$ ODEs)
   - Convergence
   - Simulation example (bifurcation analysis)
   - Kernel identification

Mixture approximations

Problem statement

$\boxed{\text{Approximate the } L^1(\mathcal{D}) \text{ function } \alpha \text{ with the following properties}}$

1. Non-negative and bounded

$$0 \leq \alpha(t) \leq K, \qquad\qquad t \in \mathcal{D}$$

2. Integrates to one

$$\int_{\mathcal{D}} \alpha(t)\,\mathrm{d}t = 1$$

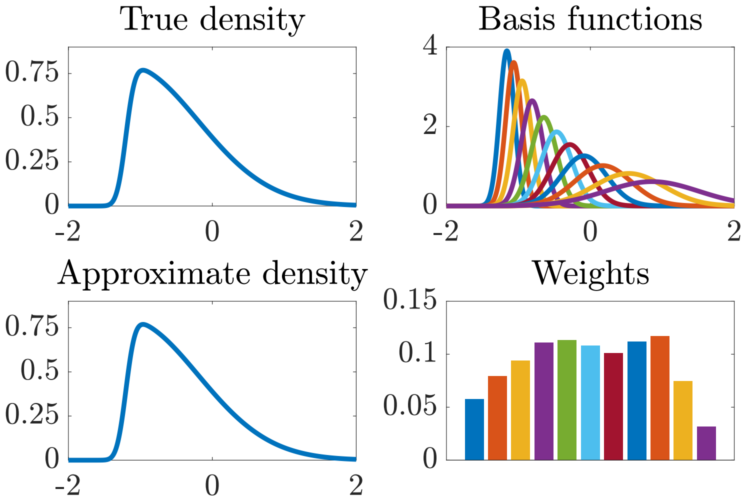
3. Continuous

The domain may be

- the real numbers, $\mathcal{D} = (-\infty, \infty)$,
- the non-negative real numbers, $\mathcal{D} = [0, \infty)$, or
- an interval, $\mathcal{D} = [a, b]$

Example for $\mathcal{D} = (-\infty, \infty)$

Skew normal distribution



Mixture approximation

Approximate α by a sum of basis functions

$$\alpha(t) \approx \hat{\alpha}(t) = \sum_{m=0}^{M} c_m \ell_m(t)$$

Weight constraints

$$\sum_{m=0}^{M} c_m = 1, \qquad 0 \leq c_m \leq 1, \qquad m = 0, \ldots, M$$

Non-negative basis functions

$$\ell_m(t) \geq 0, \qquad m = 0, \ldots, M$$

Basis functions integrate to one

$$\int_{\mathcal{D}} \ell_m(t) \, \mathrm{d}t = 1, \qquad m = 0, \ldots, M$$

Gaussian, Erlang, and beta mixture approximations

Gaussian basis functions, $t \in \mathcal{D} = (-\infty, \infty)$

$$\ell_m(t) = b_m \exp\left(-\frac{1}{2}\left(\frac{t-\mu_m}{\sigma_m}\right)^2\right), \qquad b_m = \frac{1}{\sqrt{2\pi\sigma_m^2}}$$

Erlang basis functions, $t \in \mathcal{D} = [0, \infty)$

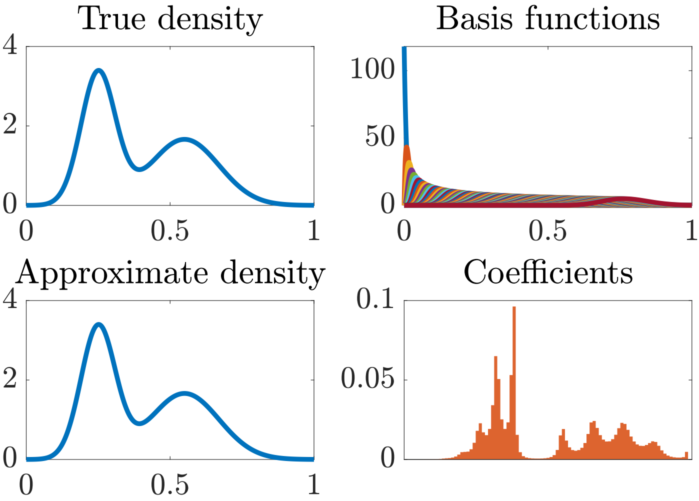
$$\ell_m(t) = b_m t^m e^{-at}, \qquad b_m = \frac{a^{m+1}}{m!}$$

Beta basis functions, $t \in \mathcal{D} = [0, \Delta t]$

$$\ell_m(t) = b_m t^m (\Delta t - t)^{M-m}, \qquad b_m = \frac{1}{\Delta t^{M+1}} \frac{(M+1)!}{m!(M-m)!}$$

Example: Erlang mixture approximation ($M = 90$)

Folded normal mixture distribution



Convergence of Erlang mixture approximations ($M = \infty$)

Coefficients

$$c_m = \int_{s_m}^{s_{m+1}} \alpha(s)\,\mathrm{d}s, \qquad s_m = m\Delta s, \qquad \Delta s = 1/a \tag{1}$$

Weak convergence (Tijms [1, 2])

$$\int_0^t \hat{\alpha}(s)\,\mathrm{d}s \to \int_0^t \alpha(s)\,\mathrm{d}s \quad \text{as} \quad a \to \infty$$

Pointwise convergence (based on delta families [3])

$$\hat{\alpha}(t) \to \alpha(t) \quad \text{as} \quad a \to \infty$$

Vitali's convergence lemma

$$\int_0^\infty |\hat{\alpha}(s) - \alpha(s)|\,\mathrm{d}s \to 0 \quad \text{as} \quad a \to \infty$$

if

- $\hat{\alpha}$ converges pointwise $\checkmark$
- $\hat{\alpha}$ is tight $\checkmark$
- $\hat{\alpha}$ is uniformly integrable $\checkmark$

Convergence of beta and Gaussian mixture approximations

1. Beta mixture approximations, $\mathcal{D} = [0, 1]$
   - ℓ_m is a Bernstein polynomial scaled by $M + 1$
   - pointwise convergence is guaranteed by Weierstrass' approximation theorem [4] as $M \to \infty$

$$c_m = \frac{1}{M+1}\alpha(s_m), \qquad s_m = m\Delta s, \qquad \Delta s = 1/M,$$
$$c_m = \int_{s_m}^{s_{m+1}} \alpha(s)\,\mathrm{d}s, \qquad s_m = m\Delta s, \qquad \Delta s = 1/(M+1)$$

2. Gaussian mixture approximations, $\mathcal{D} = (-\infty, \infty)$
   - Error analysis by Maz'ya and Schmidt [5]
   - Convergence proof by Sorenson and Alspach [6]

Optimal coefficients and rate parameter

Time domain ($k = 0, \ldots, N-1$)

$$\int_0^{t_h} \alpha(s)\,\mathrm{d}s = 1 - \epsilon, \qquad\qquad t_k = k/t_h$$

Least-squares optimization problem

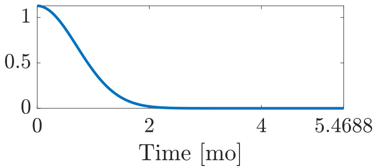
$$\begin{aligned} \min_{\{c_m\}_{m=0}^{M},a} \quad & \phi = \frac{1}{2}\sum_{k=0}^{N-1} (\alpha(t_k) - \hat{\alpha}(t_k))^2 \Delta t, \\ \text{subject to} \quad & \sum_{m=0}^{M} c_m = 1, \\ & 0 \leq c_m \leq 1, \quad m = 0, \ldots, M, \\ & a_{\min} \leq a \end{aligned}$$

- Use Matlab's `fmincon`
- Provide analytical first- and second-order derivatives

Example: Gaussian kernel

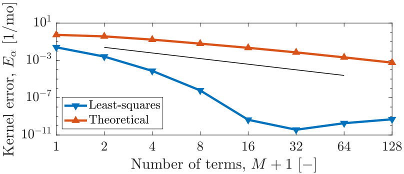
Kernel

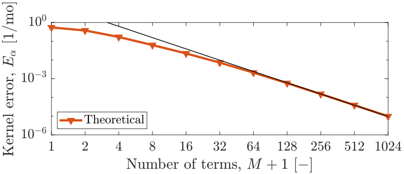
$$\alpha(t) = \frac{2}{\sqrt{\pi}} e^{-t^2}$$



Error

$$E_\alpha = \sum_{k=0}^{N_\alpha - 1} (\hat{\alpha}(t_k) - \alpha(t_k))^2 \Delta t$$





Outline

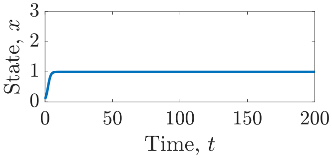
1. What are mixture approximations?
   - Three different types of mixture approximations
   - Convergence
   - Practical implementation
   - What can you do with them?
2. Mixture approximations for delay differential equations (DDEs)
   - Examples of DDEs with distributed time delays
   - The linear chain trick (DDEs $\rightarrow$ ODEs)
   - Convergence
   - Simulation example (bifurcation analysis)
   - Kernel identification

Differential equations with distributed time delays

Delay differential equations

Logistic equation

$$\dot{x}(t) = x(t)\,(1 - x(t))$$

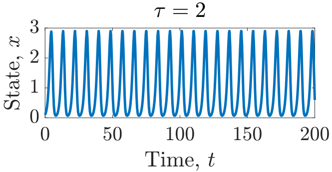


General form (ODE)

$$\dot{x}(t) = f(x(t))$$

Logistic equation w. absolute time delay

$$\dot{x}(t) = x(t)\,(1 - x(t-\tau))$$

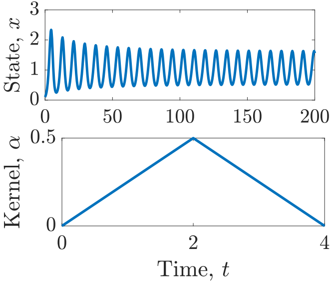


General form (DDE)

$$\dot{x}(t) = f(x(t), x(t-\tau))$$

Logistic equation w. distributed time delay

$$\dot{x}(t) = x(t)\left(1 - \int_{-\infty}^{t} \alpha(t-s)x(s)\,\mathrm{d}s\right)$$



General form (DDDE)

$$\dot{x}(t) = f(x(t), \int_{-\infty}^{t} \alpha(t-s)x(s)\,\mathrm{d}s)$$

Epidemiology

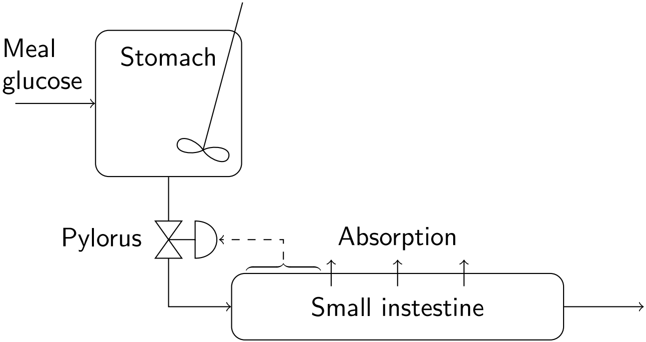
SIR model without delays

$$\begin{aligned}\dot{S}(t) &= -\beta S(t) I(t),\\ \dot{I}(t) &= \beta S(t) I(t) - \eta I(t),\\ \dot{R}(t) &= \eta I(t)\end{aligned}$$

SIR model with distributed time delay

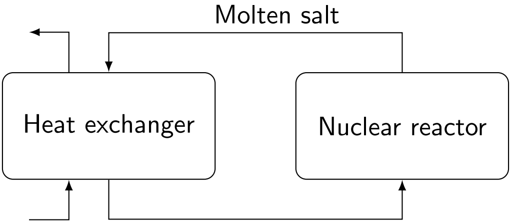
$$\begin{aligned}\dot{S}(t) &= -\beta S(t) \int_{-\infty}^{t} \alpha(t-s) I(s)\,\mathrm{d}s,\\ \dot{I}(t) &= \beta S(t) \int_{-\infty}^{t} \alpha(t-s) I(s)\,\mathrm{d}s - \eta I(t),\\ \dot{R}(t) &= \eta I(t)\end{aligned}$$

Diabetes

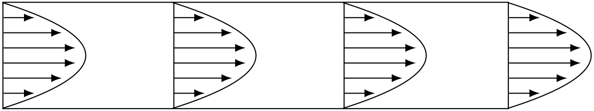


Ritschel, T.K.S., Reenberg, A.T., Carstensen, P.E., Bendsen, J., Jørgensen, J.B., 2023. Mathematical Meal Models for Simulation of Human Metabolism. arXiv: 2307.16444.

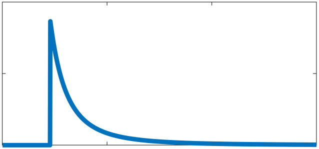
Molten salt nuclear reactor and nonuniform flow in pipes



Non-uniform velocity profile



Kernel for Hagen-Poiseuille flow (quadratic velocity profile)

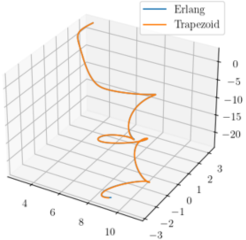


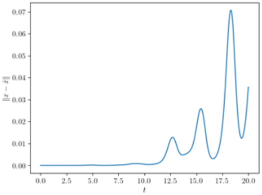
Ritschel, T.K.S., 2025. Numerical Optimal Control for Distributed Delay Differential Equations: A Simultaneous Approach based on Linearization of the Delayed Variables. In: Proceedings of the 2025 European Control Conference (ECC), June 24-27, Thessaloniki, Greece. arXiv: 2410.15083.

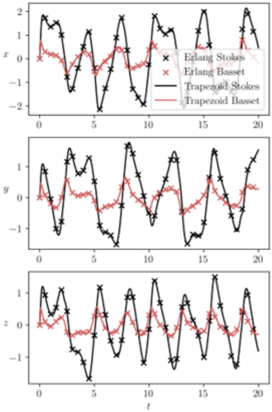
Particle flow in velocity field

Particle subject to Stoke's drag force and Basset history force

$$\dot{x}_p = u_p,$$
$$\dot{u}_p = \underbrace{\frac{1}{\mathrm{St}}\mathcal{F}(a)(u_p - u_f)}_{\text{Nonlinear drag}} + \underbrace{C\int_0^t \frac{1}{\sqrt{t-s}}(\dot{u}_p - \dot{u}_f)\,\mathrm{d}s}_{\text{Basset history force}}$$







Collaboration with PhD candidate Zejian You, Asst. Prof. Qi Wang, and Prof. Gustaaf Jacobs from San Diego State University.

Linear chain trick (DDE $\rightarrow$ ODE)

Linear chain trick

DDE with distributed time delay

$$\dot{x}(t) = f(x(t), z(t)), \qquad z(t) = \int_{-\infty}^{t} \alpha(t-s)x(s)\,\mathrm{d}s$$

Approximate kernel

$$\alpha(t) \approx \hat{\alpha}(t) = \sum_{m=0}^{M} c_m \ell_m(t)$$

Substitute into integral

$$z(t) \approx \hat{z}(t) = \sum_{m=0}^{M} c_m \int_{-\infty}^{t} \ell_m(t-s)x(s)\,\mathrm{d}s = \sum_{m=0}^{M} c_m z_m(t)$$

Differentiate z_m ($\dot{\ell}_0 = -a\ell_0(t)$, $\dot{\ell}_m(t) = a(\ell_{m-1}(t) - \ell_m(t))$)

$$\begin{aligned}\dot{z}_m(t) &= \ell_m(0)x(t) + \int_{-\infty}^{t} \dot{\ell}_m(t-s)x(s)\,\mathrm{d}s \\ &= \begin{cases} a(x(t) - z_0(t)), & m = 0, \\ a(z_{m-1}(t) - z_m(t)), & m > 0 \end{cases}\end{aligned}$$

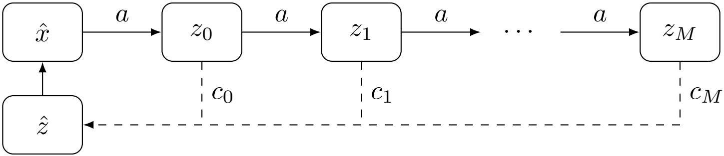
Approximate ODEs

System

$$\dot{\hat{x}}(t) = f(\hat{x}(t), \hat{z}(t)), \qquad \hat{z}(t) = \sum_{m=0}^{M} c_m z_m(t)$$

Auxiliary memory states

$$\begin{aligned} \dot{z}_0(t) &= a(\hat{x}(t) - z_0(t)), \\ \dot{z}_m(t) &= a(z_{m-1}(t) - z_m(t)), \qquad m = 1, \dots, M \end{aligned}$$



Approximate ODEs

Approximate ODEs

$$\begin{aligned}\dot{\hat{x}}(t) &= f(\hat{x}(t), \hat{z}(t)),\\ \dot{Z}(t) &= AZ(t) + B\hat{x}(t),\\ \hat{z}(t) &= CZ(t)\end{aligned}$$

Matrices

$$A = a\begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}, \qquad B = a\begin{bmatrix} 1 \\ \\ \\ \\ \end{bmatrix},$$

$$C = \begin{bmatrix} c_0 & c_1 & \cdots & c_M \end{bmatrix}$$

Does the approximate state converge ($M = \infty$)?

Bound on state error [8, Thm. 1] when f is Lipschitz and x is bounded

$$\int_0^\infty |\hat{\alpha}(t) - \alpha(t)| \, \mathrm{d}t < \epsilon \quad \Rightarrow \quad |\hat{x}(t) - x(t)| \leq \epsilon w(t) \leq \epsilon w(t_f)$$

Auxiliary function

$$\begin{aligned} w(t) &= 0, & t &\in (-\infty, t_0], \\ \dot{w}(t) &= L_z K_x + L_x w(t) + L_z \int_{-\infty}^{t} \hat{\alpha}(t-s) w(s) \, \mathrm{d}s, & t &\in [t_0, t_f] \end{aligned}$$

- The steady state of the approximate ODEs and the DDEs is the same
- The eigenvalues converge to the roots for the DDEs (Hurwitz' convergence theorem)
- Error dynamics are locally stable if the DDEs and ODEs are locally stable

Ritschel, T.K.S., 2025. On Erlang Mixture Approximations for Differential Equations with Distributed Time Delays. arXiv: 2502.12984. In submission.

Numerical example: Simulation

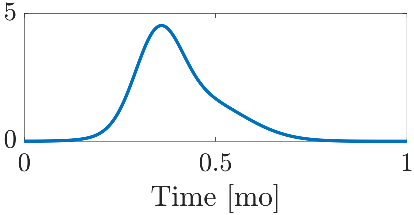
Numerical example

- System

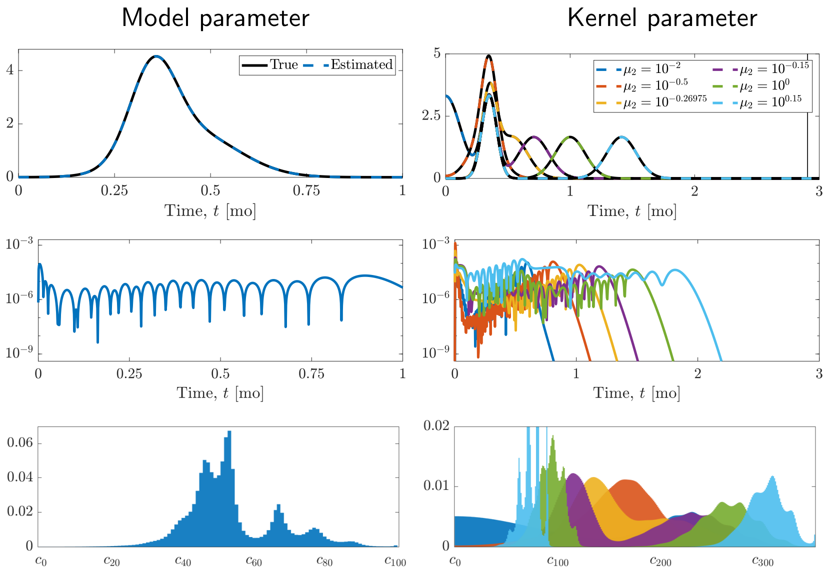
$$\dot{N}(t) = \kappa N(t)\left(1 - \frac{1}{K}\int_{-\infty}^{t} \alpha(t-s)N(s)\,\mathrm{d}s\right)$$

- Kernel

$$\alpha(t) = \gamma_1 F(t;\mu_1,\sigma_1) + \gamma_2 F(t;\mu_2,\sigma_2),$$
$$F(t;\mu,\sigma) = \frac{\exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) + \exp\left(-\frac{1}{2}\left(\frac{t+\mu}{\sigma}\right)^2\right)}{\sqrt{2\pi}\sigma}$$

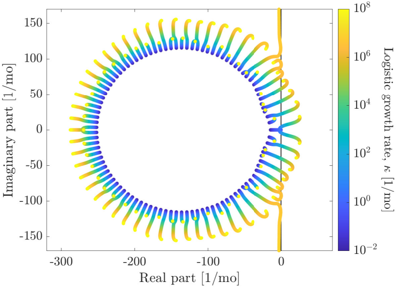


Numerical example: Bifurcation analysis

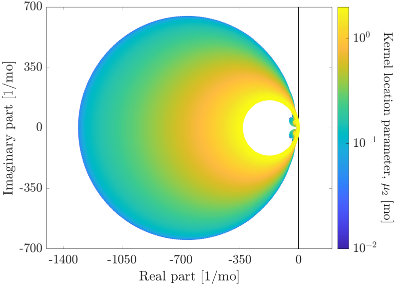


Numerical example: Bifurcation analysis

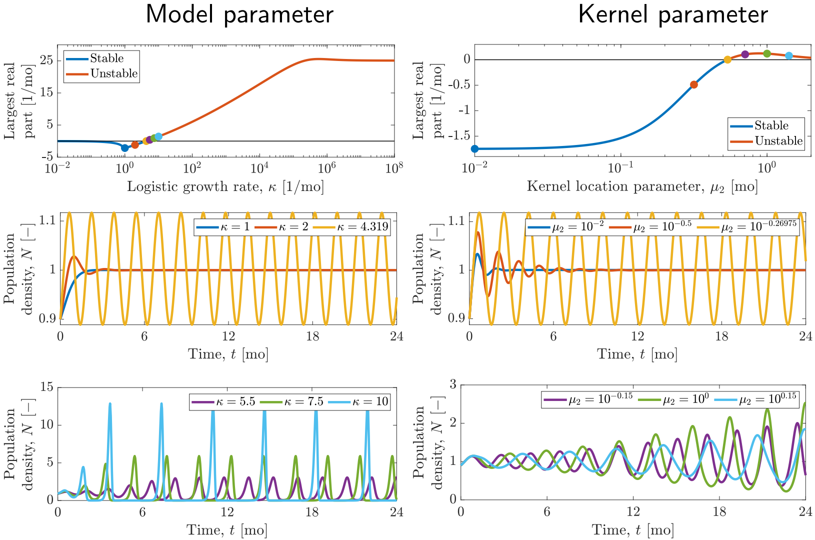
Model parameter



Kernel parameter



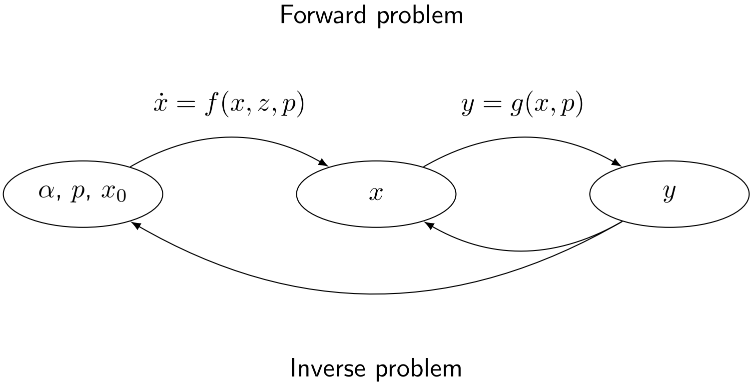
Numerical example: Bifurcation analysis



The numerical simulations are obtained with Euler's implicit method and a right rectangle rule for approximating the integral.

Kernel identification

Kernel identification (with John Wyller, NMBU, Norway)



Ritschel, T.K.S., Wyller, J., 2025. An Algorithm for Distributed Time Delay Identification without A Priori Knowledge of the Kernel. Automatica 178, pp. 112382. DOI: 10.1016/j.automatica.2025.112382.

Dynamical least-squares problem

$$\min_{\{c_m\}_{m=0}^{M}, a, p, x_0} \quad \frac{1}{2} \sum_{k=0}^{N} \|y_k - y(t_k)\|_2^2$$

subject to

$$\begin{aligned}
x(t_0) &= x_0, & & \\
Z(t_0) &= Z_0(x_0, p), & & \\
\dot{x}(t) &= f(x(t), z(t), p), & & t \in [t_0, t_f], \\
\dot{Z}(t) &= AZ(t) + Bx(t), & & t \in [t_0, t_f], \\
z(t) &= CZ(t), & & t \in [t_0, t_f], \\
y(t_k) &= g(x(t_k), p), & & k = 0, \dots, N, \\
\sum_{m=0}^{M} c_m &= 1, & & \\
a_{\min} &\leq a, & & \\
p_{\min} &\leq p \leq p_{\max}, & & \\
x_{\min} &\leq x_0 \leq x_{\max} & &
\end{aligned}$$

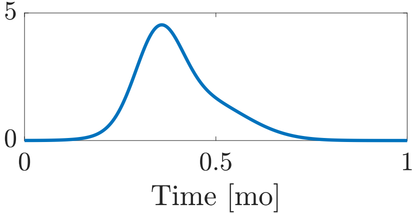
Example

- System

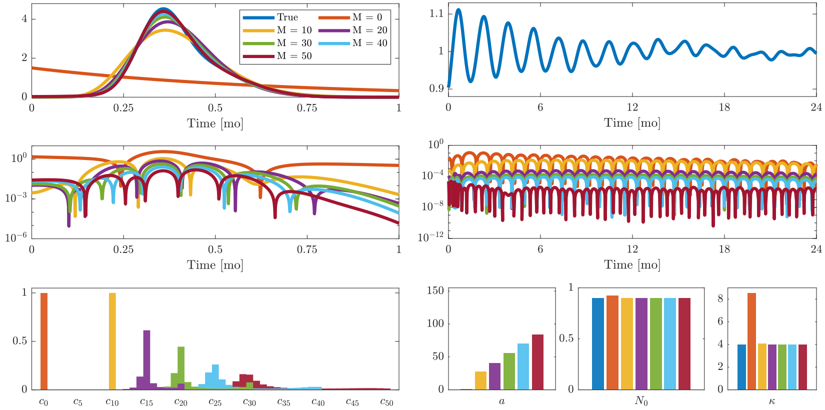
$$\dot{N}(t) = \kappa N(t)\left(1 - \frac{1}{K(t)}\int_{-\infty}^{t} \alpha(t-s)N(s)\,\mathrm{d}s\right),$$
$$K(t) = (1 + A_1\sin(2\pi\omega_1 t) + A_2\sin(2\pi\omega_2 t))\bar{K}$$

- Kernel

$$\alpha(t) = \gamma_1 F(t;\mu_1,\sigma_1) + \gamma_2 F(t;\mu_2,\sigma_2),$$
$$F(t;\mu,\sigma) = \frac{\exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) + \exp\left(-\frac{1}{2}\left(\frac{t+\mu}{\sigma}\right)^2\right)}{\sqrt{2\pi}\sigma}$$



Example



Summary

1. Approximate $L^1(\mathcal{D})$ functions

$$\alpha(t) \approx \hat{\alpha}(t) = \sum_{m=0}^{M} c_m \ell_m(t), \qquad t \in \mathcal{D}$$

2. Approximate solution to DDEs with distributed time delays

$$\dot{x}(t) = f(x(t), \int_{-\infty}^{t} \alpha(t-s) x(s) \, \mathrm{d}s)$$

3. Many different applications/possible extensions
   - Model reduction
   - Stochastic differential equations
   - Continuous thermodynamics (chemical phase/reaction equilibria)
   - Optimal/feedback/adaptive control for DDEs
   - Controllers with memory for ODEs (PID, LQR, MPC, etc.)

Bibliography I

[1] H. C. Tijms, *Stochastic models: An algorithmic approach*. Wiley Series in Probability and Statistics, Wiley, 1995.

[2] H. Cossette, D. Landriault, E. Marceau, and K. Moutanabbir, "Moment-based approximation with mixed Erlang distributions," *Variance*, vol. 10, no. 1, pp. 166–182, 2016.

[3] J. Korevaar, *Mathematical methods*, vol. 1. Academic Press, 1968.

[4] G. G. Lorentz, *Bernstein polynomials*. New York: Chelsea Publishing Company, 2nd ed., 1986.

[5] V. Maz'ya and G. Schmidt, "On approximate approximations using Gaussian kernels," *IMA Journal of Numerical Analysis*, vol. 16, pp. 13–29, 1996.

[6] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, 1971.

[7] T. K. S. Ritschel, "Numerical optimal control for distributed delay differential equations: A simultaneous approach based on linearization of the delayed variables." arXiv:2410.15083, 2024. Preprint.

Bibliography II

[8] N. Guglielmi and E. Hairer, "Applying stiff integrators for ordinary differential equations and delay differential equations to problems with distributed delays," *SIAM Journal on Scientific Computing*, vol. 47, no. 1, pp. A102–A123, 2025.

[9] T. K. S. Ritschel, "On Erlang mixture approximations for differential equations with distributed time delays." arXiv:2502.12984, 2025. In submission.

[10] T. K. S. Ritschel and J. Wyller, "An algorithm for distributed time delay identification without a priori knowledge of the kernel," *Automatica*, vol. 178, p. 112382, 2025.